



**University of
Zurich**^{UZH}

**Zurich Open Repository and
Archive**

University of Zurich
University Library
Strickhofstrasse 39
CH-8057 Zurich
www.zora.uzh.ch

Year: 2020

A Single-Stranded DNA-Encoded Chemical Library Based on a Stereoisomeric Scaffold Enables Ligand Discovery by Modular Assembly of Building Blocks

Bassi, Gabriele ; Favalli, Nicholas ; Vuk, Miriam ; Catalano, Marco ; Martinelli, Adriano ; Trenner, Anika ; Porro, Antonio ; Yang, Su ; Tham, Chuin Lean ; Moroglu, Mustafa ; Yue, Wyatt W ; Conway, Stuart J ; Vogt, Peter K ; Sartori, Alessandro A ; Scheuermann, Jörg ; Neri, Dario

Abstract: A versatile and Lipinski-compliant DNA-encoded library (DEL), comprising 366 600 glutamic acid derivatives coupled to oligonucleotides serving as amplifiable identification barcodes is designed, constructed, and characterized. The GB-DEL library, constructed in single-stranded DNA format, allows de novo identification of specific binders against several pharmaceutically relevant proteins. Moreover, hybridization of the single-stranded DEL with a set of known protein ligands of low to medium affinity coupled to a complementary DNA strand results in self-assembled selectable chemical structures, leading to the identification of affinity-matured compounds.

DOI: <https://doi.org/10.1002/adv.202001970>

Posted at the Zurich Open Repository and Archive, University of Zurich

ZORA URL: <https://doi.org/10.5167/uzh-199547>

Journal Article

Published Version



The following work is licensed under a Creative Commons: Attribution 4.0 International (CC BY 4.0) License.

Originally published at:

Bassi, Gabriele; Favalli, Nicholas; Vuk, Miriam; Catalano, Marco; Martinelli, Adriano; Trenner, Anika; Porro, Antonio; Yang, Su; Tham, Chuin Lean; Moroglu, Mustafa; Yue, Wyatt W; Conway, Stuart J; Vogt, Peter K; Sartori, Alessandro A; Scheuermann, Jörg; Neri, Dario (2020). A Single-Stranded DNA-Encoded Chemical Library Based on a Stereoisomeric Scaffold Enables Ligand Discovery by Modular Assembly of Building Blocks. *Advanced Science*, 7(22):2001970.

DOI: <https://doi.org/10.1002/adv.202001970>

A Single-Stranded DNA-Encoded Chemical Library Based on a Stereoisomeric Scaffold Enables Ligand Discovery by Modular Assembly of Building Blocks

Gabriele Bassi, Nicholas Favalli, Miriam Vuk, Marco Catalano, Adriano Martinelli, Anika Trenner, Antonio Porro, Su Yang, Chuin Lean Tham, Mustafa Moroglu, Wyatt W. Yue, Stuart J. Conway, Peter K. Vogt, Alessandro A. Sartori, Jörg Scheuermann,* and Dario Neri*

A versatile and Lipinski-compliant DNA-encoded library (DEL), comprising 366 600 glutamic acid derivatives coupled to oligonucleotides serving as amplifiable identification barcodes is designed, constructed, and characterized. The GB-DEL library, constructed in single-stranded DNA format, allows de novo identification of specific binders against several pharmaceutically relevant proteins. Moreover, hybridization of the single-stranded DEL with a set of known protein ligands of low to medium affinity coupled to a complementary DNA strand results in self-assembled selectable chemical structures, leading to the identification of affinity-matured compounds.

G. Bassi, Dr. N. Favalli, M. Vuk, M. Catalano, A. Martinelli,
Dr. J. Scheuermann, Prof. D. Neri
Department of Chemistry and Applied Biosciences
ETH Zürich
Zürich 8092, Switzerland
E-mail: joerg.scheuermann@pharma.ethz.ch;
dario.neri@pharma.ethz.ch

Dr. A. Trenner, Dr. A. Porro, Prof. A. A. Sartori
Institute of Molecular Cancer Research
University of Zürich
Zürich 8006, Switzerland

Dr. S. Yang, Prof. P. K. Vogt
Scripps Research Institute
Department of Molecular Medicine
La Jolla, CA 92037, USA

C. L. Tham, Dr. W. W. Yue
Structural Genomic Consortium (SGC)
Nuffield Department of Medicine
University of Oxford
Oxford OX1 2JD, UK

Dr. M. Moroglu, Prof. S. J. Conway
Department of Chemistry
Chemistry Research Laboratory
University of Oxford
Mansfield Road, Oxford OX1 3TA, UK

 The ORCID identification number(s) for the author(s) of this article can be found under <https://doi.org/10.1002/advs.202001970>

© 2020 The Authors. Published by Wiley-VCH GmbH. This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

DOI: 10.1002/advs.202001970

1. Introduction

The identification of small organic ligands to protein targets is crucial for many areas of chemical, biological, and pharmaceutical research. High-quality organic ligands are fundamental for the discovery of novel drugs,^[1] of affinity-chromatography supports^[2] and of tool compounds for the interrogation of biological processes.^[3]

Protein ligands can often be discovered using conventional high-throughput screening (HTS) of compound collections,

but these procedures can be expensive,^[4] require complex logistics and are typically limited to the analysis of a few hundred thousand compounds.^[5] The encoding of chemical compounds with DNA fragments serving as amplifiable identification barcodes, first postulated by Brenner and Lerner,^[6] is growing in popularity as a ligand discovery methodology, both in industry and in academia. DNA-encoded chemical libraries (DELs) do not require dedicated logistics for the handling of library members, since the DNA tag allows the facile identification of binding molecules by PCR amplification and DNA sequencing.^[7]

In most cases, DELs are interrogated by affinity-capture on target proteins of interest, immobilized on a solid support,^[8] but alternative methodologies (e.g., those based on photocrosslinking, on capillary electrophoresis or on interaction-dependent PCR procedures) have been proposed.^[9]

In principle, highly complex DELs can be constructed by the iterative assembly of multiple sets of building blocks.^[8b,10] However, the incorporation of three or more chemical moieties may result in molecular structures that violate Lipinski's rules of five and which may be too large for pharmaceutical applications.^[11] In some cases, DEL-derived binders comprising multiple building blocks may be reduced in size by retaining only those structural elements that contribute most of the binding energy.^[12] Alternatively, DELs can be constructed using two sets of building blocks, yielding "hits" which can be subsequently optimized by medicinal chemistry.^[8a,10a,13]

Here, we describe the construction of a novel versatile DEL based on two sets of building blocks, coupled to a central stereo-defined 4-azido-5-methoxy-5-oxopentanoic acid. The GB-DEL library, whose members were typically smaller than 500 Da, was

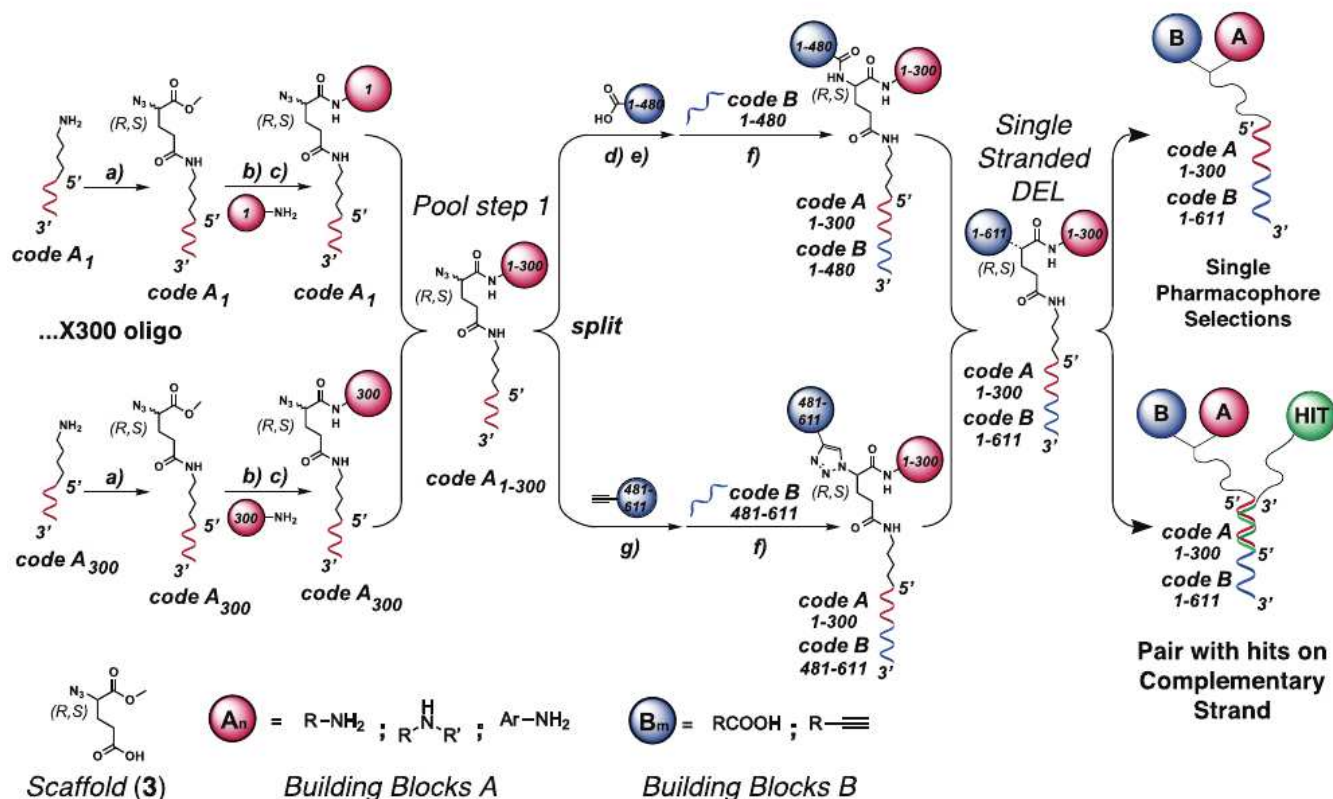


Figure 1. Schematic representation of the synthetic steps involved in GB-DEL formation. The two sets of building blocks, A and B, and the corresponding encoding DNA portions are color-coded in red and blue, respectively. The first set of building blocks A was conjugated to DNA via reverse amide bond. The building blocks B were linked to the library either by amide bond formation or CuAAC click reaction. a) Amide bond formation of three hundred different oligonucleotides (described as CodeA) with 3 (EDC/S-NHS, DIPEA, r.t. 30', in TEA buffer 50×10^{-3} M, pH = 10, 37 °C); b) deprotection of scaffold's methyl ester (LiOH 200×10^{-3} M); c) reverse amide bond formation (On-DNA-COOH, 500×10^{-3} M DMTMM in MOPS 100×10^{-3} M, 1 M NaCl, pH = 7) of each of the three hundred different scaffold-oligo conjugates with an amine (building blocks A). Each of the CodeA-building block A conjugates was individually purified by HPLC and pooled in equimolar amount to form pool step 1. The pool step 1 was split in 611 wells. d) Wells 1–480 underwent Staudinger reduction (30×10^{-3} M TCEP in 500×10^{-3} M TRIS, pH = 7.4), e) coupling to a carboxylic acid and f) encoding with CodeB 1–480 by splint ligation (Ligase buffer, NEB T-4 DNA ligase (f)); g) CuAAC click reaction (250×10^{-3} M Borate buffer, pH = 9.5, 5×10^{-3} M CuSO₄ in 5×10^{-3} M sodium ascorbate in H₂O) of N₃-pool step 1 (wells 481–611) with alkynes (10×10^{-3} M) and subsequent splint ligation (f)). After pooling and purification by RP-HPLC, the single stranded DEL was used directly in single pharmacophore selections or paired to a complementary oligonucleotide displaying a known protein binding moiety "HIT" to perform 2+1 dual pharmacophore affinity maturation selections.

screened against multiple targets, allowing the discovery of novel binders which were compatible with Lipinski's rule of five. While most DELs described so far had featured DNA in double-stranded format,^[9b,14] we used single-stranded DNA barcodes in order to enable modular lead expansion procedures by DNA self-assembly. The results of this article illustrate how novel protein ligands can efficiently be discovered in single-stranded format and how ligands can be conveniently assembled with a single-stranded DEL yielding dual-display affinity maturation libraries. This approach may practically be used in iterative assembly and selection steps allowing ligand discovery and subsequent affinity maturation.

2. Results

2.1. Library Construction

The first step for the synthesis of our library (termed GB-DEL) was performed by coupling the 4-azido-5-methoxy-5-

oxopentanoic acid (3) onto 300 distinct amino-tagged oligonucleotides (Figure 1).

We used a 1:1 mixture of the *R* and *S* stereoisomers, with the aim to assess the contribution of chirality to protein recognition at the hit validation stage. The glutamic acid derivatives were deprotected and allowed to react with 299 amines (191 primary, 36 secondary and 72 (hetero) aromatic amines), plus one code reserved for ammonia, serving as a control ("building blocks A"). The resulting conjugates were individually HPLC-purified, prior to pooling and splitting for the second coupling step with "building blocks B". At this stage, we converted 480 aliquots of the mixture (each containing the 300 building block A conjugates) into the corresponding primary amines which were later reacted with 480 carboxylic acids. In parallel, 131 aliquots of α -azido glutamic acid derivatives were reacted with terminal alkynes using copper(I)-catalyzed alkyne azide cycloaddition (CuAAC), using a previously optimized procedure.^[15] The second reactions step was encoded by splint ligation^[13c] with 611 oligonucleotides, yielding a library comprising 2 (stereoisomers) x 300 (building

block A derivatives) \times 611 (building block B derivatives) = 366 600 compounds (Figure 1 and Figures S1–S10, Supporting Information). Compounds were eventually pooled and HPLC-purified. GB-DEL in single-stranded DNA format could be used as such for protein selections (Streptavidin, CAIX, FAN1, AASS, PI3K), or hybridized to complementary oligonucleotide derivatives bearing a protein-binding moiety for affinity maturation procedures (CREBBP, HSA, CAIX, tyrosinase, and trypsin) (Figures S11–S18, Supporting Information). Specific binders with dissociation constant in the single-digit micromolar or in the submicromolar range were isolated, as illustrated in the following sections.

2.2. Selection Experiments against Protein Targets using the Library in Single-Stranded DNA Format

2.2.1. DEL Selections against Streptavidin and CAIX

Library aliquots containing 10^7 copies of each compound were used for direct library selection. In total, more than 10 000 such aliquots were available. Affinity capture experiments with proteins immobilized on magnetic beads were typically performed in triplicate, using previously described procedures.^[7a] For clarity sake, only one fingerprint is shown in the main body of the article, but replicate selection results can be found in Figures S24–S30 (Supporting Information). We characterized the unselected library by high-throughput DNA sequencing and observed a uniform distribution of sequence counts for the various library members, thus providing confidence about the proper execution of the various mixing steps during library synthesis (Figure 2a). The majority of library compounds had molecular weights <500 Da (Figure 2d) and were compliant with Lipinski's rule of five (Tables S1 and S2, Supporting Information).^[11,16]

Since most selections were performed using biotinylated protein targets, we conducted selections on “empty” streptavidin-coated magnetic beads. Only one slightly enriched compound was found which was structurally reminiscent of the high affinity streptavidin binder biotin (Figure 2e, compounds 4, 5). The compounds displayed a weak interaction with streptavidin, with dissociation constants $K_d = 37 \pm 2$ and $42 \pm 3 \times 10^{-6}$ M, respectively, which were measured by a small molecule ELISA procedure, recently described by our group.^[17] Selections were also performed against human carbonic anhydrase IX, a tumor-associated antigen and a marker of hypoxia.^[18] As expected, aromatic sulfonamides derivatives were found to be preferentially enriched (with enrichment factors (EF) up to 969-fold), as revealed by the lines at coordinates A94, A250, and B575 of the fingerprints in 3D display mode (Figure 2c). For all subsequent selections, screening results were displayed in the more compact pseudo-2D representation.^[19]

2.2.2. DEL Selections against FAN1, AASS, and PI3K

We performed selections against proteins of pharmaceutical interest (Figure 3). FANCD2-associated nuclease 1 (FAN1) is DNA repair nuclease implicated in the resolution of inter-strand crosslinks (ICLs) and in protective responses at stalled replication forks.^[20] Moreover, it has been reported that FAN1 is critical to maintain genome integrity in BRCA2-deficient cells.^[21]

Therefore, small-molecule FAN1 inhibitors may also represent a promising strategy for treatment of BRCA-mutated carriers. Based on the canSAR cancer drug discovery platform and using the chemical properties and bioactivity of small molecules annotated in ChEMBL database, it has been estimated that FAN1 is among 107 DNA damage response (DDR) factors predicted to be druggable.^[22]

Selections were performed both on His₆-tagged wild-type FAN1 and on a nuclease-dead variant (D960A) carrying a mutation in the metal-coordinating active site (Figure S31a,b, Supporting Information).^[23] A preferential enrichment of two lines, corresponding to building blocks B609 and B611 (two regioisomers of catechol derivatives) was observed (Figure 3a,b). The building block combination (A20/B611) inspired the synthesis of the two stereoisomers 6 and 7, which, however, bound to both versions of FAN1 with comparable affinities ($K_d = 0.27 \pm 0.04 \times 10^{-6}$ M; Figure 3c), most likely to a remote epitope, as they do not inhibit FAN1 enzymatic activity. Indeed, partial inhibition of FAN1 5'Flap endonuclease activity by A20/B611 was seen only in the 100×10^{-6} M range (Figure S31,c, Supporting Information).

Human α -amino adipic semialdehyde synthase (AASS) catalyzes the first two steps of lysine catabolism. Inhibition of AASS is proposed as a substrate reduction therapy^[24] for an autosomal recessive neonatal seizure disorder, pyridoxine-dependent epilepsy (PDE), caused by mutations in the ALDH7A1 gene encoding the enzyme downstream of AASS.^[25] Selection experiments against the saccharopine dehydrogenase domain of AASS revealed a preferential enrichment of the A120/B235 building block pairs (Figure 3d). Conversion of this chemical information into the stereoisomeric compounds 8 and 9 resulted in an apparent dissociation constant in the submicromolar concentration range ($K_d = 0.7 \pm 0.2 \times 10^{-6}$ M for compound 8) (Figure 3e).

The p110 α /p85 α dimeric^[26] isoform of PI3K (phosphatidylinositol 3-kinase) was the target of selections described in (Figure 3f). The regulatory subunit p85 α combines with the catalytic subunits p110 α /P85 α to form the PI3K complex. Three structurally related pairs of building blocks (A86/B371, A146/B537, and A245/B474), with nitrogen-containing heterocycles, were among the most enriched features in the fingerprints (Figure 3g). Binding assays performed with compounds 10–15 revealed a strong influence of the stereocenter on the binding affinity for the A86/B371 combination (\approx 10-fold difference between R and S stereoisomers). A lower impact of the stereocenter was observed for A146/B537 and no difference in binding affinity was seen for the two A245/B474 enantiomers (Figure 3g).

2.3. Library Expansion by 2+1 Self-Assembly with Complementary Oligonucleotide Derivatives

The single-stranded DNA format of the GB-DEL allowed the facile annealing with chemically modified complementary DNA strands, carrying a protein-binding moiety (“hit”) at the 3' end. This feature allowed us to compare selections performed with the unmodified GB-DEL and with the self-assembly products, featuring three sets of building blocks (Figure 4a). We paired the GB-DEL with a complementary DNA strand, carrying hits of intermediate affinity, in order to discover combinations of building blocks which could enable an affinity maturation procedure, using multiple proteins as targets.

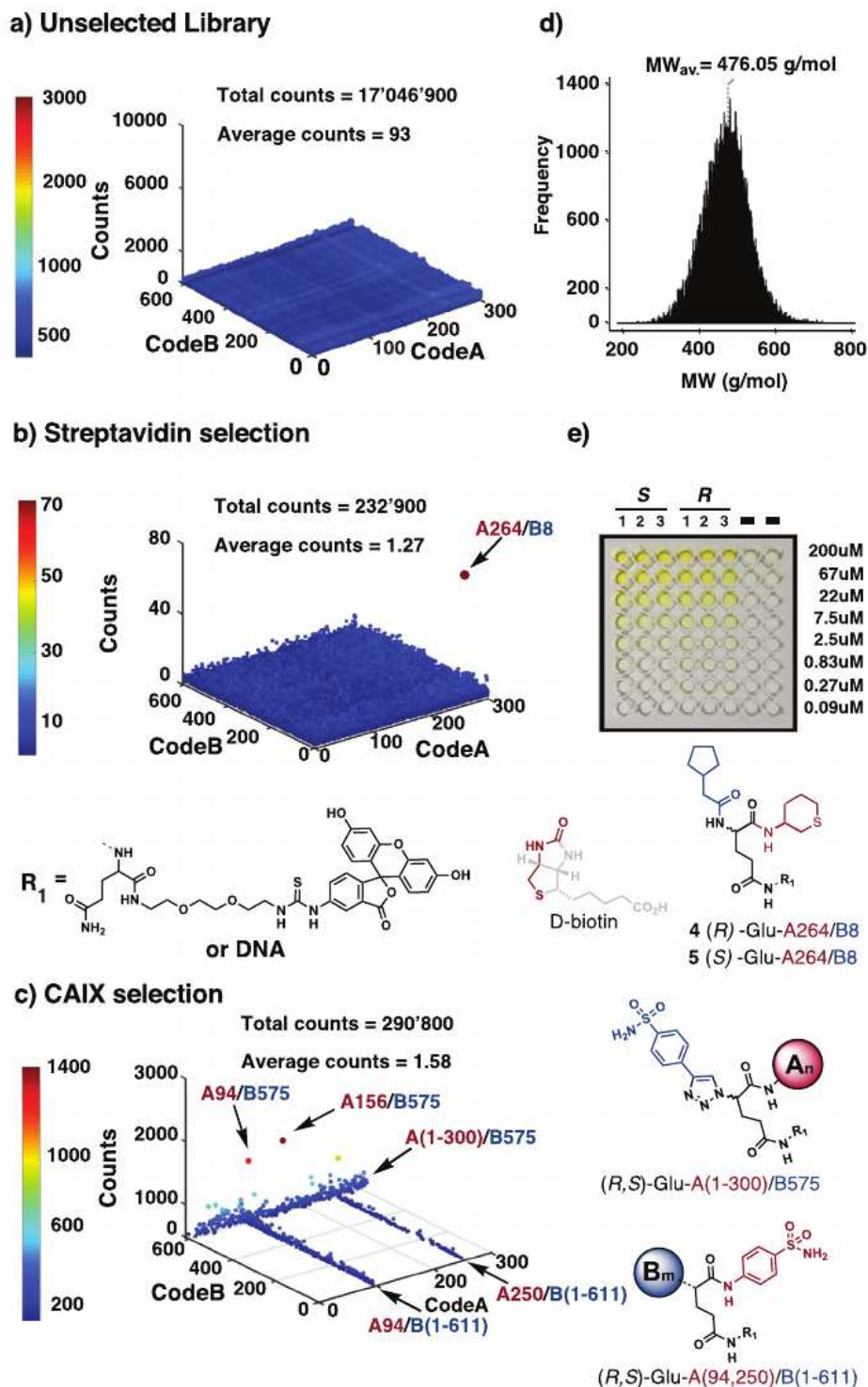
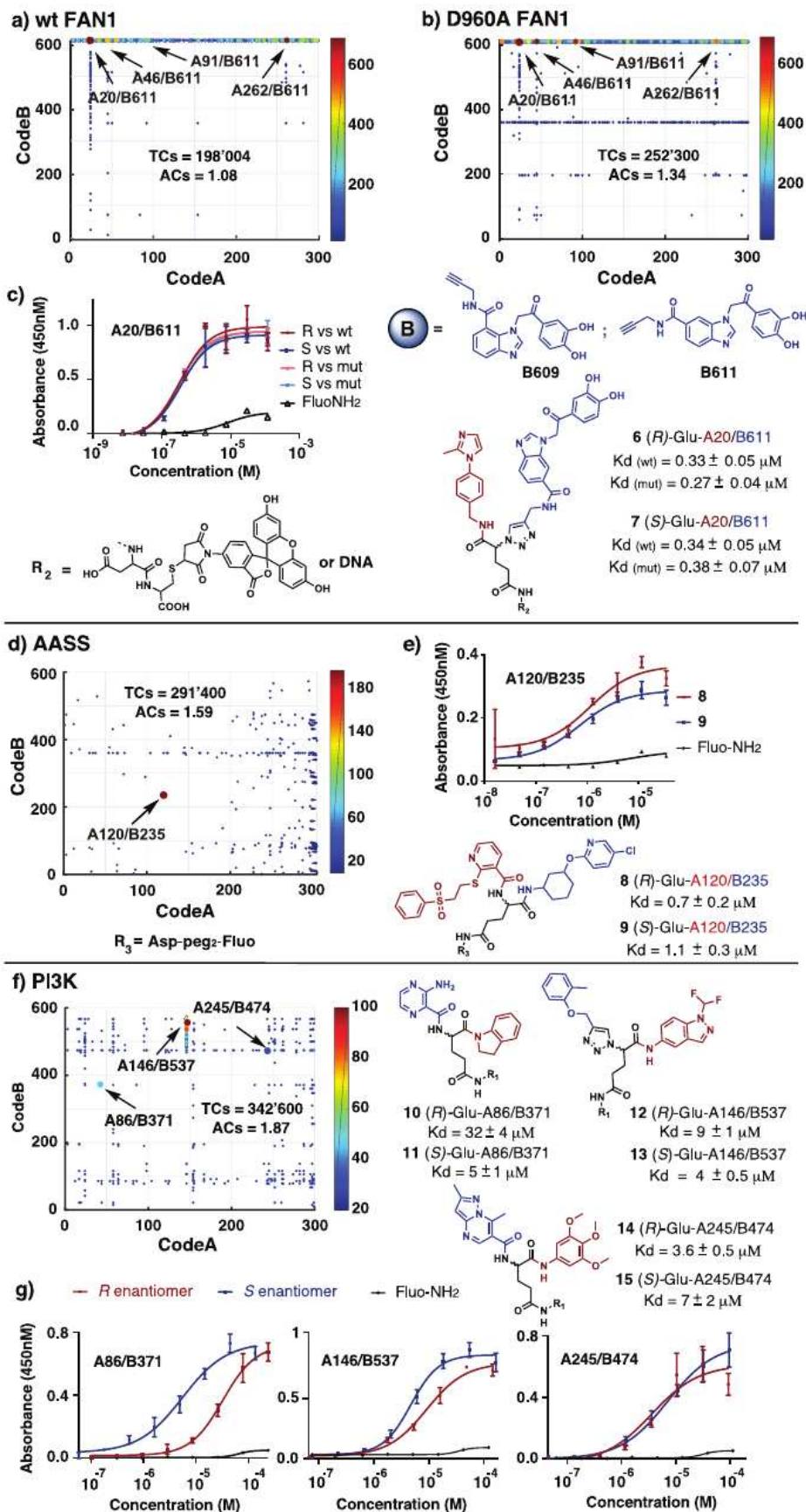


Figure 2. a) (left) High-throughput DNA sequencing (HTDS) plot of unselected library GB-DEL. The x,y,z axes of the tridimensional plot represent CodeA/CodeB and sequence counts of the GB-DEL. Total sequence counts (TCs) = 17 046 900, average sequence counts (ACs) = 93. b) HTDS plot of streptavidin selections (TCs = 232'900, ACs = 1.27). The arrow indicates the most enriched CodeA/CodeB combination (A264/B8) which was re-synthesized off-DNA (right panel compounds 4, 5). c) HTDS plot of CAIX selections. TCs = 290 800, ACs = 1.58. (Right) Aromatic sulfonamide derivatives were included in the library as positive controls: A94, A250, and B575. Indicated by the arrows are the two most enriched building block combinations, A94/B575 and A156/B575, whose molecular structures and enrichment factors can be found in Table S5 (Supporting Information). d) Histogram of GB-DEL final molecules' molecular weights distribution. On the x axis molecular weights for the CodeA/CodeB combinations, on the y axis molecular weight frequency. The mean of the Gaussian distribution corresponds to MW 476.05 g mol⁻¹. e) Enzyme-linked immunosorbent assay (ELISA) of compound 4, 5 against streptavidin. The measurements have been performed in triplicate (n = 3). The K_d values ± standard error for compound 4 and 5 are 37 ± 2 and 42 ± 3 × 10⁻⁶ M, respectively.



CREBBP is a bromodomain-containing transcriptional co-regulator, which regulates a number [27] of key intracellular processes. [28] CREBBP selections featured a benzodiazepinone (HIT1) on the complementary DNA strand. This chemical moiety had previously been described as a 32×10^{-6} M binder. [29] A preferential enrichment of the synergistic building blocks A219/B106 was observed in the selection with three sets of building blocks ("dual-pharmacophore library"), while a substantially lower enrichment was seen using the unmodified GB-DEL library ("single-pharmacophore library") (Figure 4b). Affinity measurements performed by fluorescence polarization with DNA-LNA hybrids [13c,30] revealed a binding increase for the triple combination of building blocks, while the A219/B106 moiety had an undetectable interaction with the protein at concentrations below 10×10^{-6} M. Similar findings have recently been reported for other dual-pharmacophore systems. [13c,30,31]

HSA, human serum albumin, is the most abundant protein in serum which has been used as target for pharmacokinetic modulation strategies. [32] For HSA selections, the GB-DEL was paired with a 4-(*p*-iodophenyl)butanoic acid moiety (HIT2) that we had previously described as a portable HSA binder. [13d] A striking potentiation of the A24/B137 was observed, while the A200/B44 combination had similar enrichments in both single- and dual-pharmacophore format (Figure 4c). Similar findings were observed for affinity-maturation selections against other protein targets, including the tumor-associated antigens CAIX and tyrosinase [18,33] (Figure S32, Supporting Information). In order to confirm that the binding potentiation seen "on DNA" could be translated into small organic ligands devoid of any DNA component, we resynthesized the A24/B137/HIT2 and A200/B44 HSA hits as small organic molecules 16–19 (Figure 4). The best binding combination from the "on DNA" experiments of Figure 4c also gave the best binding results "off DNA," with an apparent dissociation constant K_d for compound 19 of $56 \pm 9 \times 10^{-9}$ M (Figure 4d).

3. Conclusion

We have synthesized a novel DNA-encoded chemical library (GB-DEL), using a glutamic acid derivative as scaffold for the simultaneous display of two sets of building blocks. The library was used to isolate and characterize small molecule ligands against multiple target proteins. The strategic choice to encode the GB-DEL with DNA in single-stranded format opens the possibility to perform selection schemes based on photo-crosslinking (e.g., by hybridization with oligonucleotide-photosensitizer conjugates [9a,34] and to implement affinity-maturation with oligonucleotide-hit conjugates (Figure 4). Some of the hit compounds isolated from

the GB-DEL may be useful in practice and may deserve additional medicinal chemistry optimization.

We have discovered two novel binders of FAN1, a versatile DNA nuclease involved in ICL repair and replication stress response. The two molecules (compounds 6 and 7) bound with high affinity (K_d in the 300×10^{-9} M range) to both wt and nuclease-dead FAN1 variants (amino acids 373–1008) and exhibited only poor endonuclease inhibitory activity, suggesting that the binding site might map to a different portion of the protein. Besides the C-terminal nuclease domain, other structural regions of the FAN1 fragment used in the selections include the SAP (SAF-A/B, Acinus and PIAS) motif and the TPR (tetra-tricopeptide repeat) module responsible for homo-dimerization and DNA-binding.

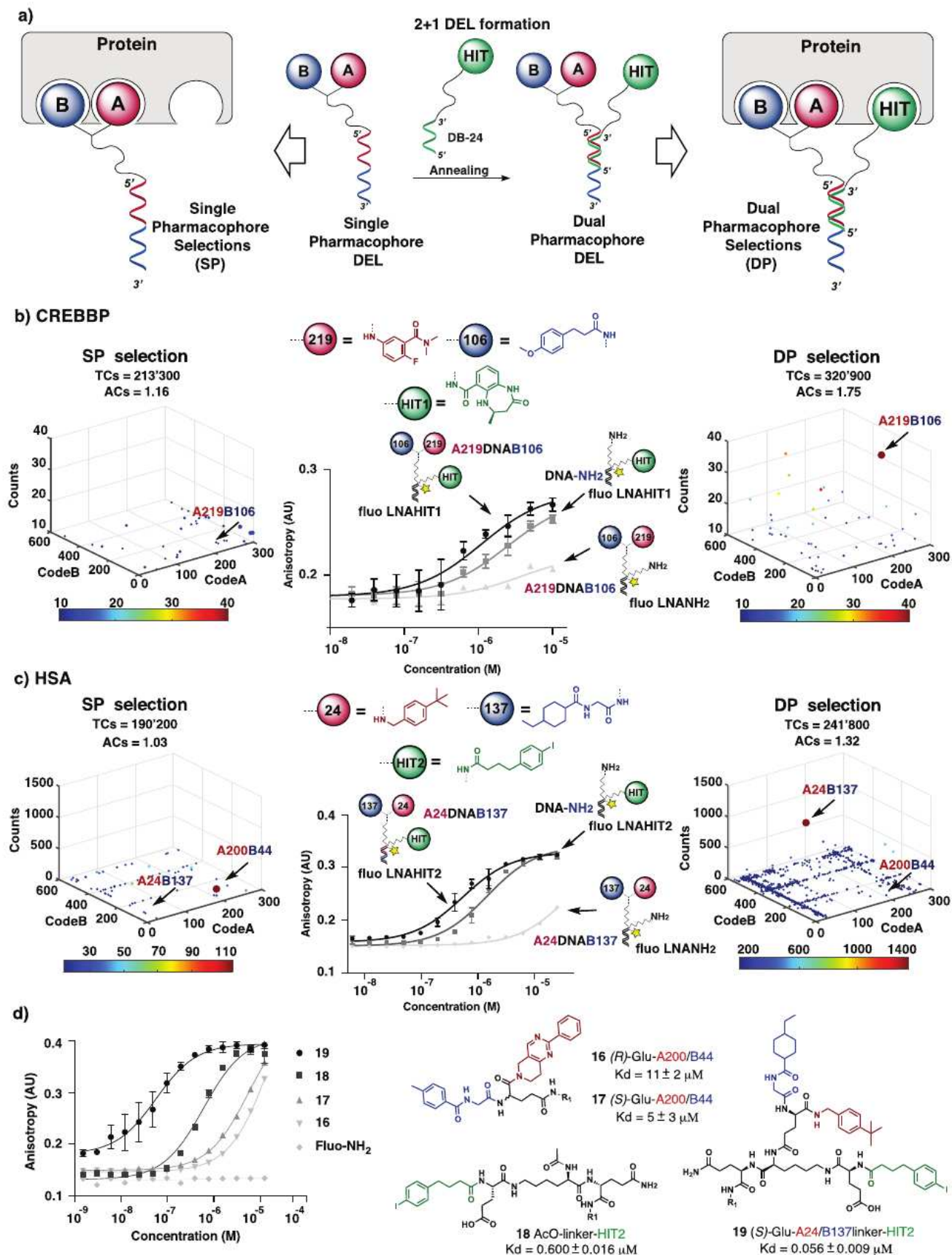
We have also isolated ligands 8 and 9 which bound tightly to AASS (alpha-aminoadipic semialdehyde synthase), an enzyme involved in the major lysine degradation pathway [35] implicated as therapy for pyridoxine-dependent epilepsy (PDE). The pathogenic driver of PDE is the accumulation of α -aminoadipate semialdehyde (AASA; a substrate for ALDH7A1) [36] and of its cyclic form L- Δ 1-piperidine-6 carboxylate (P6C). P6C reacts with pyridoxal 5'-phosphate and reduces the availability of this essential cofactor for normal cellular functions. [37] The inhibition of AASS, upstream of the ALDH7A1 pathway, may allow to reduce AASA/P6C levels and thus serve as substrate reduction therapy for the treatment of PDE. [24b]

PI3K selections have revealed an intriguing contribution of the library scaffold's chirality for protein recognition, with a ten-fold improved affinity for compound 11 over its stereoisomer 10. The protein target is a member of the lipid kinases family, which regulates signaling pathways important for cell proliferation, adhesion, survival, and motility. [38] The PIK3CA gene, coding for the p110 α subunit, is frequently mutated in cancer [39] and represents an attractive target for the development of small molecule inhibitors, some of which are currently in preclinical evaluation and some are approved therapeutics. [40]

The GB-DEL features DNA in single-stranded DNA format, which may facilitate its hybridization with chemically modified complementary oligonucleotides. This feature may be useful for a number of applications, including the execution of DEL selections with photocrosslinkers. [9a] Small ligands isolated from DELs often display suboptimal kinetic dissociation constants and may be lost during washing procedures in the screening experiments. The display of a photocrosslinker on the complementary strand may help stabilize complex formation and facilitate the recovery of preferential binders. [9c,41]

The stable non-covalent assembly of complementary DNA structures also facilitates the lead expansion of hit compounds. Our group had previously described the use of encoded

Figure 3. High-throughput DNA sequencing plots of a) wt FAN1 (TCs = 198 004, ACs = 1.08); b) D960A FAN1 (TCs = 252 500, ACs = 1.34); d) AASS (TCs = 291 400, ACs = 1.59) and f) PI3K (TCs = 342 600, ACs = 1.87). The x,y plane of the bidimensional plots represent Code A and Code B, respectively. The jet scale and the dots color code indicate code sequence counts. The arrows indicate Code A/Code B combinations selected for re-synthesis (compounds 6–15), the colors, red and blue, indicate building block A and B, respectively. c,e,g) ELISA binding assays of compounds against their protein targets. c) Compounds 6–7 against wt and D960A FAN1; e) compounds 8–9 against AASS; g) compounds 10–15 against PI3K; on the x axis concentration (M), on the y axis absorbance at 450 nm. Red and blue curves indicate R and S stereoisomer, respectively. K_d measurements have been performed in triplicate ($n = 3$). Error bars on the curve indicate standard deviation of the three measurements. The K_d values \pm standard error for compound 6–15 are given next to the chemical structures. Structures of CodeA/CodeB combinations selectively enriched in the FAN1 selection (A46/B611, A91/B611, and A262/B611) can be found in Table S5 (Supporting Information).



self-assembling chemical (ESAC) libraries as an avenue for the discovery of synergistic pairs of building blocks that may interact with the cognate target in a bidentate fashion and provide an affinity gain through the “chelate effect”.^[13c,42] The affinity maturation strategy described in this article represents an extension of the ESAC strategy, leading to the identification of three chemical moieties which may synergistically interact with the target protein of choice. The procedure was exemplified for the affinity maturation of ligands for the CREBBP bromodomain and HSA. Compound re-synthesis in the absence of DNA, combined with a judicious linker optimization, may lead to an additional gain in binding affinity (e.g., $K_d = 56 \pm 9 \times 10^{-9}$ M for the HSA binder). Indeed, the molecular features of the linker may positively or negatively affect the affinity constant of the resulting bidentate ligand.^[30]

In spite of its relatively small size (366 600 compounds), the GB-DEL generated hits against multiple target proteins of pharmaceutical interest. The glutamic acid scaffold provided a directional display of functional group pairs which could synergistically interact with the cognate protein. The reactions used for library synthesis proceeded with excellent conversion yields,^[10a] contributing to the overall quality and reproducibility of screening experiments which were conducted in triplicate. The single-stranded DNA format used for library encoding is likely to find a broader applicability in the future, not only because of its versatility but also because it often yields better selection results compared to the double-stranded DNA format in terms of ligand recovery and signal-to-noise ratios.^[43]

Supporting Information

Supporting Information is available from the Wiley Online Library or from the author.

Acknowledgements

Financial support by the ETH Zürich, the Swiss National Science Foundation (grant number 310030_182003/1) and by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement 670603) is gratefully acknowledged. The National Cancer Institute (R35 CA197582 to P.K.V.) is gratefully acknowledged. (This is TSRI ms. #29981.) M.M. and S.J.C. thank the EP-SRC Centre for Doctoral Training in Synthesis for Biology and Medicine (EP/L015838/1) and AstraZeneca, Diamond Light Source, Defense Science and Technology Laboratory, Evotec, GlaxoSmithKline, Janssen, Novartis, Pfizer, Syngenta, Takeda, UCB, and Vertex for studentship support.

Work in the Sartori laboratory is supported by the Swiss National Science Foundation (grant number: 31003A_176161) and the Swiss Cancer League (grant number: KFS-4702-02-2019). The Structural Genomics Consortium is a registered charity (Number 1097737) that receives funds from AbbVie, Bayer Pharma AG, Boehringer Ingelheim, Canada Foundation for Innovation, Eshelman Institute for Innovation, Genome Canada, Innovative Medicines Initiative (EU/EFPIA) [ULTRA-DD grant no. 115766], Janssen, Merck & Co., Novartis Pharma AG, Ontario Ministry of Economic Development and Innovation, Pfizer, São Paulo Research Foundation-FAPESP, Takeda, and Wellcome Trust [092809/Z/10/Z]. The authors warmly thank Mr. Sebastian Oehler for preliminary studies and fruitful discussions. The authors also thank Jonathan Held for his help on the synthesis of compound 8 and 9. The authors are also grateful to the Functional Genomics Center Zurich for help with high-throughput DNA sequencing. Instant JChem (ChemAxon) was used for structure and data management (<http://www.chemaxon.com>).

Conflict of Interest

D.N. is co-founder and shareholder of Philochem AG (<http://www.philochem.com>), a company active in the field of DNA-encoded chemical libraries.

Keywords

Human alpha-aminoadipic semialdehyde synthase (AASS), affinity maturation, cyclic-AMP response element binding protein (CREBBP), DNA-encoded chemical libraries (DEL), drug discovery, encoded self-assembling chemical libraries (ESAC), FANCD2-associated nuclease 1 (FAN-1), phosphatidylinositol 3-kinase (PI3K)

Received: May 26, 2020

Revised: August 17, 2020

Published online: October 14, 2020

- [1] a) T. E. Nielsen, S. L. Schreiber, *Angew. Chem., Int. Ed. Engl.* **2008**, 47, 48; b) D. A. Erlanson, S. W. Fesik, R. E. Hubbard, W. Jahnke, H. Jhoti, *Nat. Rev. Drug Discovery* **2016**, 15, 605; c) R. H. Folmer, *Drug Discovery Today* **2016**, 21, 491.
- [2] M. Bergseid, A. R. Baytan, J. P. Wiley, W. M. Ankener, M. L. Stolowitz, K. A. Hughes, J. D. Chesnut, *BioTechniques* **2000**, 29, 1126.
- [3] a) A. Mullard, *Nat. Rev. Drug Discovery* **2019**, 18, 733; b) C. H. Arrow-smith, J. E. Audia, C. Austin, J. Baell, J. Bennett, J. Blagg, C. Bountra, P. E. Brennan, P. J. Brown, M. E. Bunnage, C. Buser-Doepner, R. M. Campbell, A. J. Carter, P. Cohen, R. A. Copeland, B. Cravatt, J.

Figure 4. a) Schematic representation of 2+1 DEL formation (middle): The single pharmacophore DEL (GB-DEL) is annealed to a partially complementary oligonucleotide (DB-24) bearing a protein-binding moiety (“HIT”). Left and right parts represent the selection process in single pharmacophore (SP) and in dual pharmacophore 2+1 (DP) mode, respectively. b) High-throughput DNA sequencing plots (HTDS) of selections against the CREBBP bromodomain in SP format (TCs = 213 300, ACs = 1.16, left graph) and DP format bearing DB-24-HIT1 (TCs = 320 900, ACs = 1.75, right graph). The x,y,z axes of the tridimensional plot represent CodeA/CodeB and sequence counts of the GB-DEL. The arrows indicate the A109/B206 combination (Enrichment factor (EF) SP selection = 3, EF DP selection = 22) selected for affinity maturation of HIT1. (Middle): On-DNA fluorescence polarization (FP) of affinity matured hit against CREBBP: A219, B106, HIT1 displayed on 12-mer DNA/fluor-LNA hetero-duplex (black curve), only HIT1 displayed on 12-mer DNA/fluor-LNA hetero-duplex (dark grey) and A219/B106 displayed on 12-mer DNA/fluor-LNA hetero-duplex (light grey). c) HTDS plots of selections against human serum albumin in SP format (TCs = 190 200 ACs = 1.03, left graph) and DP format bearing DB-24-HIT2 (TCs = 241 800, ACs = 1.32, right graph). The x,y,z axes of the tridimensional plot represent CodeA/CodeB and sequence counts of the GB-DEL. The arrows indicate the combination A24/B137 (Enrichment factor (EF) SP selection = 3, EF DP selection = 1119) selected for affinity maturation and A200/ B244 combination selected for de novo discovery EF SP selection = 112, respectively. EF has been calculated according to Equations S1–S3 in Tables S5 and S6 (Supporting Information). (Middle): On-DNA fluorescence polarization (FP) with HSA: A24/B137/ HIT2 displayed on 12-mer DNA/fluor-LNA hetero-duplex (black curve), only HIT2 displayed on 12-mer DNA/fluor-LNA hetero-duplex (dark grey) and only A24/B137 (light grey). d) Off-DNA fluorescence polarization of HSA hits. All the K_d measurements have been performed in triplicate ($n = 3$). Error bars on the curve indicate standard deviation of the three measurements. The K_d values \pm standard error for compound 16–19 are given next to the chemical structures.

- L. Dahlin, D. Dhanak, A. M. Edwards, M. Frederiksen, S. V. Frye, N. Gray, C. E. Grimshaw, D. Hepworth, T. Howe, K. V. Huber, J. Jin, S. Knapp, J. D. Kotz, R. G. Kruger, D. Lowe, M. M. Mader, B. Marsden, A. Mueller-Fahrnow, S. Muller, R. C. O'Hagan, J. P. Overington, D. R. Owen, S. H. Rosenberg, B. Roth, R. Ross, M. Schapira, S. L. Schreiber, B. Shoichet, M. Sundstrom, G. Superti-Furga, J. Taunton, L. Toledo-Sherman, C. Walpole, M. A. Walters, T. M. Willson, P. Workman, R. N. Young, W. J. Zuercher, *Nat. Chem. Biol.* **2015**, *11*, 536; c) A. C. Dar, K. M. Shokat, *Annu. Rev. Biochem.* **2011**, *80*, 769.
- [4] R. A. Goodnow Jr., C. E. Dumelin, A. D. Keefe, *Nat. Rev. Drug Discovery* **2017**, *16*, 131.
- [5] A. A. Shelat, R. K. Guy, *Nat. Chem. Biol.* **2007**, *3*, 442.
- [6] S. Brenner, R. A. Lerner, *Proc. Natl. Acad. Sci. USA* **1992**, *89*, 5381.
- [7] a) W. Decurtins, M. Wichert, R. M. Franzini, F. Buller, M. A. Stravs, Y. Zhang, D. Neri, J. Scheuermann, *Nat. Protoc.* **2016**, *11*, 764; b) L. Mannocci, Y. Zhang, J. Scheuermann, M. Leimbacher, G. De Bellis, E. Rizzi, C. Dumelin, S. Melkko, D. Neri, *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 17670; c) F. Buller, M. Steiner, J. Scheuermann, L. Mannocci, I. Nissen, M. Kohler, C. Beisel, D. Neri, *Bioorg. Med. Chem. Lett.* **2010**, *20*, 4188; d) G. Zhao, Y. Huang, Y. Zhou, Y. Li, X. Li, *Expert Opin. Drug Discovery* **2019**, *14*, 735.
- [8] a) M. Leimbacher, Y. Zhang, L. Mannocci, M. Stravs, T. Geppert, J. Scheuermann, G. Schneider, D. Neri, *Chemistry* **2012**, *18*, 7729; b) N. Favalli, S. Biendl, M. Hartmann, J. Piazza, F. Sladojevich, S. Gräslund, P. J. Brown, K. Näreoja, H. Schuler, J. Scheuermann, R. Franzini, D. Neri, *ChemMedChem* **2018**, *13*, 5; c) M. A. Clark, R. A. Acharya, C. C. Arico-Muendel, S. L. Belyanskaya, D. R. Benjamin, N. R. Carlson, P. A. Centrella, C. H. Chiu, S. P. Creaser, J. W. Cuzzo, C. P. Davie, Y. Ding, G. J. Franklin, K. D. Franzen, M. L. Geftter, S. P. Hale, N. J. Hansen, D. I. Israel, J. Jiang, M. J. Kavarana, M. S. Kelley, C. S. Kollmann, F. Li, K. Lind, S. Mataruse, P. F. Medeiros, J. A. Messer, P. Myers, H. O'Keefe, M. C. Oliff, C. E. Rise, A. L. Satz, S. R. Skinner, J. L. Svendsen, L. Tang, K. van Vloten, R. W. Wagner, G. Yao, B. Zhao, B. A. Morgan, *Nat. Chem. Biol.* **2009**, *5*, 647.
- [9] a) G. Li, Y. Liu, Y. Liu, L. Chen, S. Wu, Y. Liu, X. Li, *Angew. Chem., Int. Ed. Engl.* **2013**, *52*, 9544; b) D.-Y. Wang, Y. Cao, L.-Y. Zheng, L.-D. Chen, X.-F. Chen, Z.-Y. Hong, Z.-Y. Zhu, X. Li, Y.-F. Chai, *Chem. - Eur. J.* **2017**, *23*, 9; c) P. Zhao, Z. Chen, Y. Li, D. Sun, Y. Gao, Y. Huang, X. Li, *Angew. Chem., Int. Ed. Engl.* **2014**, *53*, 10056; d) L. M. McGregor, D. J. Gorin, C. E. Dumelin, D. R. Liu, *J. Am. Chem. Soc.* **2010**, *132*, 15522; e) L. M. McGregor, T. Jain, D. R. Liu, *J. Am. Chem. Soc.* **2014**, *136*, 3264.
- [10] a) Y. Li, R. De Luca, S. Cazzamalli, F. Pretto, D. Bajic, J. Scheuermann, D. Neri, *Nat. Chem.* **2018**, *10*, 441; b) B. A. Seigal, W. H. Connors, A. Fraley, R. M. Borzilleri, P. H. Carter, S. L. Emanuel, J. Fargnoli, K. Kim, M. Lei, J. G. Naglich, M. E. Pokross, S. L. Posy, H. Shen, N. Surti, R. Talbott, Y. Zhang, N. K. Terrett, *J. Med. Chem.* **2015**, *58*, 2855.
- [11] L. Z. Benet, C. M. Hosey, O. Ursu, T. I. Oprea, *Adv. Drug Delivery Rev.* **2016**, *101*, 89.
- [12] a) S. B. Berger, P. Harris, R. Nagilla, V. Kasparcova, S. Hoffman, B. Swift, L. Dare, M. Schaeffer, C. Capriotti, M. Ouellette, B. W. King, D. Wisnoski, J. Cox, M. Reilly, R. W. Marquis, J. Bertin, P. J. Gough, *Cell Death Discovery* **2015**, *1*, 15009; b) J. P. Maiani, A. McFedries, Z. H. Foda, R. E. Kleiner, X. Q. Du, M. A. Leissring, W. J. Tang, M. J. Charron, M. A. Seeliger, A. Saghatelian, D. R. Liu, *Nature* **2014**, *511*, 94.
- [13] a) F. Samain, T. Ekblad, G. Mikutis, N. Zhong, M. Zimmermann, A. Nauer, D. Bajic, W. Decurtins, J. Scheuermann, P. J. Brown, J. Hall, S. Gräslund, H. Schuler, D. Neri, R. M. Franzini, *J. Med. Chem.* **2015**, *58*, 5143; b) R. M. Franzini, T. Ekblad, N. Zhong, M. Wichert, W. Decurtins, A. Nauer, M. Zimmermann, F. Samain, J. Scheuermann, P. J. Brown, J. Hall, S. Gräslund, H. Schuler, D. Neri, *Angew. Chem., Int. Ed. Engl.* **2015**, *54*, 3927; c) M. Wichert, N. Krall, W. Decurtins, R. M. Franzini, F. Pretto, P. Schneider, D. Neri, J. Scheuermann, *Nat. Chem.* **2015**, *7*, 241; d) C. E. Dumelin, S. Trussell, F. Buller, E. Trachsel, F. Bootz, Y. Zhang, L. Mannocci, S. C. Beck, M. Drumea-Mirancea, M. W. Seeliger, C. Baltes, T. Muggler, F. Kranz, M. Rudin, S. Melkko, J. Scheuermann, D. Neri, *Angew. Chem., Int. Ed. Engl.* **2008**, *47*, 3196; e) J. Scheuermann, C. E. Dumelin, S. Melkko, Y. Zhang, L. Mannocci, M. Jaggi, J. Sobek, D. Neri, *Bioconjugate Chem.* **2008**, *19*, 778.
- [14] a) L. H. Yuen, S. Dana, Y. Liu, S. I. Bloom, A. G. Thorsell, D. Neri, A. J. Donato, D. Kireev, H. Schuler, R. M. Franzini, *J. Am. Chem. Soc.* **2019**, *141*, 5169; b) C. J. Gerry, M. J. Wawer, P. A. Clemons, S. L. Schreiber, *J. Am. Chem. Soc.* **2019**, *141*, 10225.
- [15] N. Favalli, G. Bassi, T. Zanetti, J. Scheuermann, D. Neri, *Helv. Chim. Acta* **2019**, *102*, 1900033.
- [16] C. A. Lipinski, *Drug Discovery Today: Technol.* **2004**, *1*, 337.
- [17] L. Prati, M. Bigatti, E. J. Donckele, D. Neri, F. Samain, *BBRC* **2020**, unpublished.
- [18] C. T. Supuran, *Nat. Rev. Drug Discovery* **2008**, *7*, 168.
- [19] D. Neri, R. Lerner, *Annu. Rev. Biochem.* **2018**, *87*, 479.
- [20] a) K. Kratz, B. Schopf, S. Kaden, A. Sandoel, R. Eberhard, C. Lademann, E. Cannavo, A. A. Sartori, M. O. Hengartner, J. Jiricny, *Cell* **2010**, *142*, 77; b) C. Lachaud, A. Moreno, F. Marchesi, R. Toth, J. J. Blow, J. Rouse, *Science* **2016**, *351*, 846.
- [21] A. Porro, M. Berti, J. Pizzolato, S. Bologna, S. Kaden, A. Saxer, Y. Ma, K. Nagasawa, A. A. Sartori, J. Jiricny, *Nat. Commun.* **2017**, *8*, 1073.
- [22] L. H. Pearl, A. C. Schierz, S. E. Ward, B. Al-Lazikani, F. M. Pearl, *Nat. Rev. Cancer* **2015**, *15*, 166.
- [23] A. Smogorzewska, R. Desetty, T. T. Saito, M. Schlabach, F. P. Lach, M. E. Sowa, A. B. Clark, T. A. Kunkel, J. W. Harper, M. P. Colaiacovo, S. J. Elledge, *Mol. Cell* **2010**, *39*, 36.
- [24] a) J. Kopec, I. Pena, E. Rembeza, M. McLaughlin, O. Fedorov, C. Strain-Damerell, S. Goubin, S. MacKinnon, N. Burgess-Brown, P. E. Brennan, A. MacKenzie, P. Arruda, W. W. Yue, <https://doi.org/10.5281/zenodo.1219674>; b) W. W. Yue, S. Mackinnon, G. A. Bezerra, *Emerging Top. Life Sci.* **2019**, *3*, 63.
- [25] I. A. Pena, L. A. Marques, A. B. Laranjeira, J. A. Yunes, M. N. Eberlin, A. MacKenzie, P. Arruda, *Biochim. Biophys. Acta, Mol. Basis Dis.* **2017**, *1863*, 121.
- [26] C. H. Huang, D. Mandelker, O. Schmidt-Kittler, Y. Samuels, V. E. Velculescu, K. W. Kinzler, B. Vogelstein, S. B. Gabelli, L. M. Amzel, *Science* **2007**, *318*, 1744.
- [27] M. Brand, J. Clayton, M. Moroglu, M. Schiedel, S. Picaud, J. P. Bluck, A. Skwarska, A. K. N. Chan, C. M. C. Laurin, A. R. Scorch, K. F. L. See, T. P. C. Rooney, O. Fedorov, G. Perell, P. Kalra, W. A. Cortopassi, K. E. Christensen, R. I. Cooper, R. S. Paton, W. C. K. Pomerantz, P. C. Biggin, E. M. Hammond, P. Filippakopoulos, S. J. Conway, *ChemRxiv* **2020**, unpublished.
- [28] a) M. E. Breen, A. K. Mapp, *Curr. Opin. Chem. Biol.* **2018**, *45*, 195; b) T. P. Rooney, P. Filippakopoulos, O. Fedorov, S. Picaud, W. A. Cortopassi, D. A. Hay, S. Martin, A. Tumber, C. M. Rogers, M. Philpott, M. Wang, A. L. Thompson, T. D. Heightman, D. C. Pryde, A. Cook, R. S. Paton, S. Muller, S. Knapp, P. E. Brennan, S. J. Conway, *Angew. Chem., Int. Ed. Engl.* **2014**, *53*, 6126; c) M. Schiedel, S. J. Conway, *Curr. Opin. Chem. Biol.* **2018**, *45*, 166.
- [29] A. M. Taylor, A. Cote, M. C. Hewitt, R. Pastor, Y. Leblanc, C. G. Nasveschuk, F. A. Romero, T. D. Crawford, N. Cantone, H. Jayaram, J. Setser, J. Murray, M. H. Beresini, G. de Leon Boenig, Z. Chen, A. R. Conery, R. T. Cummings, L. A. Dakin, E. M. Flynn, O. W. Huang, S. Kaufman, P. J. Keller, J. R. Kiefer, T. Lai, Y. Li, J. Liao, W. Liu, H. Lu, E. Pardo, V. Tsui, J. Wang, Y. Wang, Z. Xu, F. Yan, D. Yu, L. Zawadzke, X. Zhu, X. Zhu, R. J. Sims 3rd, A. G. Cochran, S. Bellon, J. E. Audia, S. Magnuson, B. K. Albrecht, *ACS Med. Chem. Lett.* **2016**, *7*, 531.
- [30] M. Bigatti, A. Dal Corso, S. Vanetti, S. Cazzamalli, U. Rieder, J. Scheuermann, D. Neri, F. Sladojevich, *ChemMedChem* **2017**, *12*, 1748.

- [31] A. Dal Corso, M. Catalano, A. Schmid, J. Scheuermann, D. Neri, *Angew. Chem., Int. Ed. Engl.* **2018**, 57, 17178.
- [32] A. Zorzi, S. Linciano, A. Angelini, *MedChemComm* **2019**, 10, 1068.
- [33] M. B. Bloom, D. Perry-Lalley, P. F. Robbins, Y. Li, M. el-Gamil, S. A. Rosenberg, J. C. Yang, *J. Exp. Med.* **1997**, 185, 453.
- [34] A. Sannino, A. Gironda-Martinez, E. M. D. Gorre, L. Prati, J. Piazza, J. Scheuermann, D. Neri, E. J. Donckele, F. Samain, *ACS Comb. Sci.* **2020**, 22, 204.
- [35] A. Hallen, J. F. Jamie, A. J. Cooper, *Amino Acids* **2013**, 45, 1249.
- [36] P. B. Mills, E. Struys, C. Jakobs, B. Plecko, P. Baxter, M. Baumgartner, M. A. Willemsen, H. Omran, U. Tacke, B. Uhlenberg, B. Weschke, P. T. Clayton, *Nat. Med.* **2006**, 12, 307.
- [37] a) C. Brocker, M. Cantore, P. Failli, V. Vasiliou, *Chem. -Biol. Interact.* **2011**, 191, 269; b) G. Scharer, C. Brocker, V. Vasiliou, G. Creadon-Swindell, R. C. Gallagher, E. Spector, J. L. Van Hove, *J. Inherited Metab. Dis.* **2010**, 33, 571.
- [38] a) B. Vanhaesebroeck, L. Stephens, P. Hawkins, *Nat. Rev. Mol. Cell Biol.* **2012**, 13, 195; b) N. P. Shanware, K. Bray, R. T. Abraham, *Annu. Rev. Pharmacol. Toxicol.* **2013**, 53, 89.
- [39] F. Janku, J. J. Wheler, A. Naing, G. S. Falchook, D. S. Hong, V. M. Stepanek, S. Fu, S. A. Piha-Paul, J. J. Lee, R. Luthra, A. M. Tsimberidou, R. Kurzrock, *Cancer Res.* **2013**, 73, 276.
- [40] X. Wang, J. Ding, L. H. Meng, *Acta Pharmacol. Sin.* **2015**, 36, 1170.
- [41] A. Sannino, E. Gabriele, M. Bigatti, S. Mulatto, J. Piazza, J. Scheuermann, D. Neri, E. J. Donckele, F. Samain, *ChemBioChem* **2019**, 20, 955.
- [42] S. Melkko, J. Scheuermann, C. E. Dumelin, D. Neri, *Nat. Biotechnol.* **2004**, 22, 568.
- [43] G. Bassi, N. Favalli, S. Oehler, A. Martinelli, M. Catalano, J. Scheuermann, D. Neri, *Biochem. Biophys. Res. Commun.* **2020** May 5;S0006-291X(20)30761-0. <https://doi.org/10.1016/j.bbrc.2020.04.035>. On-line ahead of print.